

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-126446

(43)Date of publication of application : 15.05.1998

(51)Int.Cl. H04L 12/56
H04L 29/08

(21)Application number : 09-234517

(71)Applicant : LUCENT TECHNOL INC

(22)Date of filing : 29.08.1997

(72)Inventor : LAKSHMAN T V
MADHOW UPAMANYU
SUTER BERNHARD

(30)Priority

Priority number : 96 25122 Priority date : 30.08.1996 Priority country : US
97 858310 19.05.1997

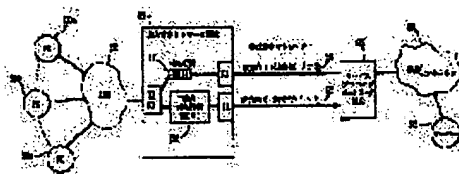
US

(54) DATA TERMINAL IN TCP NETWORK

(57)Abstract:

PROBLEM TO BE SOLVED: To obtain a data terminal and a data transmission controller in a TCP/IP network that transfers a data packet by deciding that a receiving confirmation packet and a data packet are sent from an upper-stream direction receiving confirmation queue and an upper-stream direction data queue.

SOLUTION: A queue 97 of a subscriber network terminal 100 temporarily stores a received data packet and distributes it to a subscriber's LAN through a specified route. A client terminal (PC) sends a receiving confirmation (ACK) packet, and a service provider network terminal 85 receives it through a backward (upstream) direction link 95. The client PC sends a data packet to the terminal 100 by multiplexing data to a LAN 55. A composite queue 150 of the terminal 100 forms each queue in every connection and transfers both a data packet and an ACK packet to the terminal 85 via the link 95.



LEGAL STATUS

[Date of request for examination] 05.04.1999

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 3321043

[Date of registration] 21.06.2002

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of

BEST AVAILABLE COPY

rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-126446

(43) 公開日 平成10年(1998) 5月15日

(51) Int.Cl.⁶

H 0 4 L 12/56
29/08

識別記号

F I

H 0 4 L 11/20
13/00

1 0 2 A
3 0 7 Z

審査請求 未請求 請求項の数11 OL (全 9 頁)

(21) 出願番号 特願平9-234517

(22) 出願日 平成9年(1997) 8月29日

(31) 優先権主張番号 60/025122

(32) 優先日 1996年8月30日

(33) 優先権主張国 米国 (US)

(31) 優先権主張番号 08/858310

(32) 優先日 1997年5月19日

(33) 優先権主張国 米国 (US)

(71) 出願人 596077259

ルーセント テクノロジーズ インコーポ
レイテッド

Lucent Technologies
Inc.

アメリカ合衆国 07974 ニュージャージ
ー、マレーヒル、マウンテン アベニュー
600-700

(72) 発明者 ティ. ヴィ. ラクシュマン

アメリカ合衆国、07724 ニュージャージ
ー、イートンタウン、ピクトリア ドライ
ブ 118

(74) 代理人 弁理士 三俣 弘文

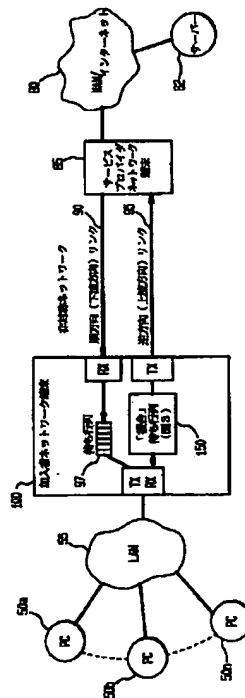
最終頁に続く

(54) 【発明の名称】 TCPネットワーク内のデータ端末

(57) 【要約】

【課題】 データパケットを転送するTCP/IPネットワーク内のデータ端末とデータ伝送制御装置を提供する。

【解決手段】 (a) 逆方向リンク上の輻輳点で例えばバッファで以前に受信したACKの代わりの最後のACKを保存し、(b) 十分に管理された方法で関連メモリ(バッファ)を含む上流方向リンクのバンド幅を割り当て、(c) 重み付けが下流方向リンクに関連したパラメータに基づいて行われるよう上流方向リンクのバンド幅を割り当てるように、公平待ち行列系あるいは重み付けラウンドロビン系を用いることである。これは各接続の順方向パス対逆方向パスのバンド幅の比率を所定の境界値以下に維持しながら行う。



【特許請求の範囲】

【請求項1】 (A) 下流方向リンク(90)を介して、ソースからデータパケットを受信し、前記下流方向データパケットを下流方向待ち行列(97)に記憶するプロセッサ(101)と、前記プロセッサ(101)は、前記上流方向待ち行列が満杯状態であることに一致して新たな受領確認パケットの生成に応答し、

(B) 前記下流方向データパケットの受信に応答して生成される受領確認パケットを記憶し、上流方向受領確認待ち行列中に既に記憶されている最も古い受領確認パケットの代わりに、前記上流方向受領確認待ち行列内に新たに生成された受領確認パケットを記憶する上流方向受領確認待ち行列と、

(C) 上流方向リンク(95)を介して伝送されるべきデータパケットを記憶する上流方向データ待ち行列と、

からなるTCPネットワーク内のデータ端末(100)において、

前記上流方向受領確認待ち行列と前記データ待ち行列とは、前記下流方向リンク(90)に関連した伝送パラメータの関数として決定される関連重み付け値を有し、前記重み付け値は、前記上流方向リンク(95)を介して伝送するためにそれぞれ前記上流方向受領確認待ち行列と、上流方向データ待ち行列から受領確認パケットとデータパケットとを送信するよう決定することを特徴とするTCPネットワーク内のデータ端末(100)。

【請求項2】 前記下流方向リンク(90)に関連づけられた伝送パラメータは、下流方向伝送レート C_{down} (ビット/秒)と下流方向データバッファ B_{down} のサイズを含むことを特徴とする請求項1のデータ端末。

【請求項3】 サービスプロバイダからこれと共通のネットワーク端末(100)を介して接続された加入者端末(50)への情報パケットの通信を可能にする下流方向リンク(90)と、前記加入者端末(50)から前記ネットワーク端末(100)を介して前記サービスプロバイダへパケットを伝送する上流方向リンク(95)と、を有する非対称通信システム用のデータ伝送制御装置において、

(A) 前記ネットワーク端末(100)に対し、前記上流方向リンク(95)を介して伝送するために前記ネットワーク端末(100)により受信された情報パケットを分類する分類手段と、前記情報はデータパケットと受領確認パケットとを有し、

(B) 送信されるべき情報パケットの分類に応じて前記ネットワーク端末(100)内の上流方向リンク接続待ち行列を形成する制御手段(101)と、前記上流方向接続待ち行列(151)は、データ情報伝送用のデータ待ち行列と、前記下流方向リンク(90)を介して伝送されたデータパケットに応答してACKパケットを送信するACK待ち行列とを有し、前記制御手段(10

1)は、ある上流方向リンク接続待ち行列内に記憶された最も古いACKパケットをその後受信したACKパケットで置換し、

(C) 前記上流方向接続待ち行列(151)の各々にサービスし、前記上流方向リンク接続待ち行列から前記パケットを上流方向リンク(95)と下流方向リンク

(90)の両方の伝送パラメータに応じて前記サービスプロバイダに送信するスケジューラ(300)とを有することを特徴とするデータ伝送制御装置。

10 【請求項4】 前記ネットワーク端末(100)は、前記上流方向接続待ち行列(151)が形成されるメモリ記憶手段を有し、

前記制御手段(101)は、加入者端末(50)とサービスプロバイダとの間を接続するために、個別の上流方向リンク接続待ち行列を形成するために前記メモリ記憶を割り当てることを特徴とする請求項3の装置。

【請求項5】 前記分類手段は、形成された上流方向リンク接続待ち行列と、その関連するサービスプロバイダ及び加入者端末との接続のテーブルを生成する手段を有し、

前記制御手段(101)は、前記分類手段に応じて関連サービスプロバイダと加入者端末接続状態に応じて受信したパケットを上流方向接続待ち行列(151)に入力することを特徴とする請求項3の装置。

【請求項6】 前記ネットワーク端末(100)は、前記サービスプロバイダと前記加入者端末(50)との間の前記下流方向リンク(90)を介して送信されたパケットを一時的に記憶する下流方向バッファを有することを特徴とする請求項3の装置。

30 【請求項7】 前記スケジューラ(300)は、重み付け系に応じて前記上流方向リンク接続待ち行列にサービスし、

前記上流方向リンク接続待ち行列は、そのサイズと占有率に応じて優先重み付け値が割り当てられ、この重み付け値に応じて前記上流方向リンク(95)を介して前記パケットを送信することを特徴とする請求項3の装置。

【請求項8】 前記スケジューラ(300)は、公平待ち行列系に応じて前記上流方向リンク接続待ち行列にサービスし、

40 前記上流方向リンク接続待ち行列は、その等しい重み値が割り当てられ、この重み値に応じて前記上流方向リンク(95)を介して前記パケットを送信することを特徴とする請求項3の装置。

【請求項9】 前記下流方向リンク(90)に関連づけられた伝送パラメータは、下流方向伝送レート C_{down} (ビット/秒)と下流方向データバッファ B_{down} のサイズを含むことを特徴とする請求項3の装置。

【請求項10】 前記上流方向リンク(95)に関連づけられた伝送パラメータは、上流方向伝送レート C_{up} (ビット/秒)とACKパケットの長さ L_{ACK} を含む

ことを特徴とする請求項3の装置。

【請求項11】 前記ネットワーク端末(100)は、TCP/IP接続を含むことを特徴とする請求項3の装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、ネットワーク接続を介してデータパケットを転送するプロトコルに関し、特にTCP/IPと称するインターネットトランスポートプロトコルを介してデータを転送する方法に関する。

【0002】

【従来の技術】近い将来、住宅地の加入者に対する高速インターネットアクセスサービスが非対称アクセスネットワークを用いて提供することが予想されている。この非対称アクセスネットワークは、例えば(a)非対称デジタル加入者ライン(Asymmetric Digital Subscriber Line(ADSL))を採用するネットワークあるいはその変形例と、(b)ハイブリッドファイバ同軸(Hybrid Fiber Coax(HFC))ネットワークと、(c)下流方向バス(ネットワークから加入者へ)はケーブルリンクであり、上流方向バス(加入者からネットワークあるいはサービスプロバイダへ)は電話線であるようなネットワークを含む。

【0003】このようなシステムの特徴は、ある種のケーブルモデムあるいはADSLアクセスサービスが、ネットワークで用いられた場合には、10倍程度のバンド幅の非対称性あるいはリターンバスが電話線である場合には、100倍ものバンド幅の非対称性を有することである。このためこの非対称のネットワークの特徴は、ネットワークから加入者へは高速の下流方向リンクを有し、加入者からネットワークへは低速の上流方向リンクを有することである。そしてこのようなネットワークは、データを転送するのに公知のインターネットトランスポートプロトコルTCP/IPを用いている。

【0004】簡単に説明すると、TCP/IPは、ウィンドウベースのフロー制御プロトコルであり、データソースは、順番に番号のついたデータパケットを宛先に転送する。そして宛先は、データパケットの受領に応じていわゆる蓄積型の受領確認(cumulative acknowledgment(ACK))を返送する。そしてこの蓄積型受領確認通知は、シーケンス番号により次の予測されるデータパケットを特定する。

【0005】宛先がシーケンス番号Nまでの全てのデータパケットを正しく受信した場合には、宛先は、次に受領しようとしているデータパケットはパケット番号がN+1であることを示す。TCP/IPにおいては、データソースは動的に変化するウィンドウサイズを保持する。W個のパケットのウィンドウサイズとは、ソースが宛先から受領した受領確認通知がN+1であることを示

した場合には、n+Wまでのシーケンス番号を有するデータパケットを送信できることを意味する。

【0006】ソースと宛先との間の往復遅延が大きい場合には、ウィンドウサイズは使用バンド幅を効率よく管理するために通常大きな値である。このことはTCPにおいては伝送されたデータパケットが失われるまでウィンドウサイズを増加し、そしてその後このウィンドウサイズをより効率的な値まで減少することにより行われる。ウィンドウサイズを減少する量は、個別形態のTCP(例えばTCP-TahoeとTCP-Reno)毎に異なる。

【0007】TCP/IPにおけるパケットの喪失は、「高速再転送」スキームを用いて検出される。この高速再転送スキームは、同一の次の予測データパケット(例、N+1)をそれぞれが識別する数個の受領確認(以下、ACKSと称する)を受領することに基づいている。このようなACKSは、N+1番目のパケットが失われ、再転送する必要があることを示す。

【0008】「高速再転送」系は、あるパケット(例、n+1)が失われ、複数の一連のパケットを宛先が正しく受信した場合に発動される。しかし、この方法はn+1番目以降の全てのパケットが失われたか、あるいはn+1番目のパケットがソースにより送信された最後のパケットであるである場合には高速再転送パケット系は、喪失を検出できない。

【0009】このような可能性を処理するために、ある種の検出システムは、例えば500ミリ秒毎のような粗い動きをする「粗い」タイマーの時間経過を用いてこのようなパケットの喪失を検出している。すなわちパケットが転送されるとタイマーがスタートして対応するACKの受領を待つ。ACKをタイマーの時間経過の満了前に受領しない場合には、パケットは喪失したものと見なされる。これにより、伝送リンクはその期間アイドル状態となる。新たなパケットは通常タイマーの時間経過の前に送信されることはないからである。

【0010】上述したことから明らかなように、TCPの接続は、ネットワークの両方向に設定される。特に、従来のシングルTCP接続を示す図1のように順方向接続は、高速下流方向リンク25を用いて送信者15(ネットワーク内の)からパケットをデータ端末16(加入者側)に搬送する。一方、データ端末16は低速上流方向リンク26を用いてACKsを送信者15に送る。

【0011】逆に、逆方向接続は、低速上流方向リンク26を用いてデータ端末16からのデータパケットを送信者15に送り、かつ高速下流方向リンク25を用いてネットワーク送信者15からの受領確認通知(ACK)をデータ端末16に送る。TCPの宛先が成功裏に受信した各データパケットに対し、ACKを戻す。多くのACKは、明示的なものであるが、TCPは輻輳制御用にACKを用いているかである。

【0012】高速下流方向リンク25を介して受領した大量のデータパケットにより、同じく大量のACKが低速上流方向リンク26を介して搬送される。このことは明らかに上流方向リンク26に輻輳を生ぜしめ、そして、バッファ30では逆方向接続用のデータパケットの喪失につながる。上述したように、このような喪失は逆方向接続に対しては、それぞれのウィンドウのサイズを減少させることになり、ひいては上流方向リンクの資源（例、バッファとバンド幅）が、これらの接続に制御しながら割り当てられない限りスループットの減少を引き起こす。

【0013】同時にまた、ACKの喪失は、送信者側のバースチネス（burstiness）の増加を引き起こす。例えば、ウィンドウサイズをWと仮定してさらにソースがパケットnに対するACKを待っていると仮定すると、ウィンドウが完全に使用された場合には、n+Wまでが送信される。さらにまた宛先が全てのパケットを成功裏に受信したが、nからn+4までのパケットのACKは喪失し、n+5のパケットに対するACKがソースにより成功裏に受信されたと仮定する。

【0014】するとソースが後者（n+5パケット）のACKを受信すると、ソースは次のデータパケットはn+6のパケットであることが分かり、そのためn+5+Wまでのパケットの送信を許可する。ソースは、n+Wまでのパケットを送信することができるだけであるために次の5個のパケットを送信することが許される。

【0015】そしてバッファが5に満たない場合には、このことはパケットが失われる原因となり、対応するウィンドウサイズの減少を引き起こす。同時に全ての問題となっているACKパケットの喪失は、必然的にタイムアウトとなりウィンドウサイズを減少させ、その結果システム効率を減少させる。Wのサイズが小さい場合（例えばW=1の場合）には、常に発生する可能性がある。

【0016】

【発明が解決しようとする課題】したがって本発明の目的は、データパケットを転送するTCP/IPネットワーク内のデータ端末とデータ伝送制御装置を提供することである。

【0017】

【課題を解決するための手段】このため前述したことにより、本発明者は、逆方向接続の適正な動作のために上流方向リンク上で十分なバンド幅を確保する必要がある、そしてこのようなバンド幅は下流方向リンクを介してデータパケットを受信した結果、ACKを転送するために保持する必要があると認識した。

【0018】ACKSの伝送を取り扱うために十分なバンド幅が割り当てられない場合には、順方向接続の順方向パスと逆方向パスの間のバンド幅の非対称により、順方向パス上で過剰なバースチネス（excessive burstiness）が発生し、これがデータパケットの喪失につなが

り、その結果ウィンドウサイズが減少しそれに起因してスループットが減少することになる。このため上流方向リンク上でバンド幅を確保することは、順方向接続および逆方向接続の両方にとって良好な性能を維持するために十分管理する必要がある。

【0019】これは本発明によれば、（a）逆方向リンク上の輻輳点で例えばバッファで以前に受信したACKの代わりに最後のACKを保存し、（b）十分に管理された方法で関連メモリ（バッファ）を含む上流方向リンクのバンド幅を割り当て、（c）重み付けが下流方向リンクに関連したパラメータに基づいて行われるよう上流方向リンクのバンド幅を割り当てるように、公平待ち行列系（fair-queuing scheme）あるいは重み付けラウンドロビン系（weighted round-robin scheme）を用いることである。これは各接続の順方向パスと逆方向パスのバンド幅の比率を所定の境界値以下に維持しながら行う。

【0020】具体的に説明すると、輻輳が発生すると宛先から送信された最後（最も最近）のACKを保存する。これは例えば最後のACKの保存用に空きを形成するために輻輳した上流方向バッファの「フロント/トップ」に記憶されたACKを脱落させることにより行う。同時に逆方向接続上のデータパケットの喪失は、データパケットの転送用に上流方向リンク上にバンド幅を保存することにより防ぐことができる。

【0021】本発明の一側面によれば、このようなバンド幅の保存は、戻り（上流方向）パス上にデータパケット用とACK用に論理的に別個のデータバッファを用いることにより行われる。さらに上流方向リンクのサービスの実行は、接続あたりの待ち行列と共に重み付けラウンドロビンサービスに基づいて行われ、本発明の一側面によればこのような重み付けは、下流方向リンクの特定のパラメータに従って選択される。

【0022】特に例えば、下流方向リンクが毎秒 C_d 個のパケットの伝送比率を有し、上流方向リンク上のバッファのサイズが B_d パケットの場合には、順方向接続用の上流方向リンク上のACK用に利用できるバンド幅は、少なくとも毎秒 C_d / B_d のACKでなければならない。上流方向リンクの残りのバンド幅の全ては、上流方向リンクを介して伝送される逆接続からのデータパケットに割り当てられる。この個々のバッファと重み付けサービス系は、逆方向接続がバンド幅の公平な分配を得るのに必要である。

【0023】

【発明の実施の形態】図2はTCPネットワーク接続のデータフロー制御装置を表し、同図は本発明のTCP接続は複数のユーザPC50a, 50b, . . . , 50nにサービスしている。図2に示すように各ユーザである顧客PC50a, . . . , 50nは、LAN55に接続され、このLAN55は接続すなわち加入者ネットワーク

端末100を有する。この加入者ネットワーク端末100は、順方向（下流方向）リンク90からデータパケットを受領し、例えばWAN/インターネット80によりサービスされるWAN/インターネット80からのサービスプロバイダからデータパケットを受信する。

【0024】具体的に説明すると、これらのデータパケットはサービスプロバイダネットワーク端末85から順方向（下流方向）リンク90を介して通信され、これをLAN55が受領し、そしてこのLAN55はこのパケットを分配すなわち分離して適切な顧客端末（PC）に分配配送する。加入者ネットワーク端末100は待ち行列97を有し、この待ち行列97は受信したデータパケットを一時的に蓄積し、ある特定のレートでもって加入者LANに分配する。顧客PCは受領確認（ACK）パケットを送信し、これをサービスプロバイダネットワーク端末85が逆方向（上流方向）リンク95を介して受領する。

【0025】顧客PCはデータパケットをLAN55に送信し、そしてこのLAN55はデータを多重化して加入者ネットワーク端末100に送信する。前述したように従来技術に係るTCP接続アーキテクチャの特徴は、サービスプロバイダネットワーク端末85からLAN55への順方向（下流方向）リンク90と、LAN55からサービスプロバイダネットワーク端末85への逆方向（上流方向）リンク95とを有する非対称構成で、これによりボトルネックと、パケットのランダム喪失と、受け入れ難い遅延と、低スループットを引き起こす。

【0026】パケットの最大スループットを確保するために加入者ネットワーク端末100は、本発明の複合待ち行列150を具備し、この複合待ち行列150は接続毎の個々の待ち行列を形成し、データパケットとACKパケットの両方をサービスプロバイダネットワーク端末85に逆方向（上流方向）リンク95を介して適宜転送する。インターネット通信ではTCP-TahoeとTCP-Renoのプロトコルが、それぞれ順方向リンクと逆方向リンクを介した上流方向ビットレートと下流方向ビットレートとを規定している。

【0027】図3に示すように複合待ち行列150は、メモリ蓄積領域を有し、複数の接続毎の待ち行列151a, . . . , 151nの設定を容易にしている。データスループットを最大にするために図3に示すように複合待ち行列150は、フロントエンドプロセッサ101として構成されており、このフロントエンドプロセッサ101は待ち行列管理機能（メモリと待ち行列マネージャ160として示されている）を実行し、パケット分離機能（クラシーファイヤ（分類器）200）とパケット伝送スケジューリング機能（スケジューラ300）を実行する。

【0028】このフロントエンドプロセッサ101により実行される上記の機能により、複数のバッファ待ち行

列はメモリと待ち行列管理スケジュール計画に従って逆方向（上流方向）リンク95へ、そして逆方向リンク95上の上流方向の情報パケットを記憶し伝送するよう必要に応じて形成され形成される。

【0029】例えばフロントエンドプロセッサ101の制御下では、クラシーファイヤ200は加入者ネットワークの各来入パケットを検査して、その種類例えばデータパケットか、ACK等を分類する。そしてクラシーファイヤ200は既存の接続の追跡を行い、そしてその種類（グリーディ、散在、データ、ACK）を追跡し、それぞれの待ち行列151a, . . . , 151nを各接続に割り当て、そしてパケットを正しい待ち行列に配置する。

【0030】具体的に説明すると、各ソースPC50a, b, cは対応するネットワークアドレスとそれに関連したポートとを有し、データまたはACKを宛先アドレスとポートに送信するときにはダイナミックな待ち行列接続がメモリ内に設定され、宛先アドレス/ポートへ最終的に送信されるよう情報を記憶する。クラシーファイヤ200は特定のソースアドレス/ポート用とその対応する宛先アドレス/ポート用に形成した待ち行列接続のテーブル（図示せず）を生成して維持し、特定のソースと宛先との間のデータあるいはACKの上流方向通信の追跡を行う。

【0031】システムのコストおよび/またはメモリ条件によって1つの接続毎の待ち行列151が、例えばデータパケットの受領用にダイナミックに形成され、また別の待ち行列は送信ソースの数に拘らず、ACKパケットの受領用に形成される。しかし、待ち行列は各個別のTCP接続毎に形成されるのが好ましい。実際には待ち行列の数は制限され、そして待ち行列の占有スキームはメモリと待ち行列マネージャにより強制的に全ての待ち行列が占有されたときにはダイナミックに置換されるようにする。

【0032】図3に示すように、メモリと待ち行列マネージャ160はパケットをメモリ待ち行列に記憶して待ち行列占有方策を強制する。特に全ての待ち行列が占有され、そして新たなソース/宛先接続が形成されると、待ち行列マネージャはダイナミックに待ち行列を新たなソース/宛先接続間で新たなパケットを上流方向に伝送できるように利用可能にする。

【0033】実行中のTCP/IPプロトコル、例えばTCP-Tahoeの制限内では、あるスキームは時間切れ機能を実行し、これにより待ち行列マネージャは、別のソース/宛先接続用に形成された待ち行列の実行不可能性を決定する。即ち、占有された待ち行列は所定の時間内（例えば5分以内）ではパケットの受信も行わないし送信も行わない。クラシーファイヤ200内の行列あたりの接続のテーブルはそれに応じて更新される。

【0034】メモリと待ち行列マネージャ160は第1

受信パケットルールを実行するが、このルールは、待ち行列マネージャはいかなる環境下でも最後の受信パケットを待ち行列に常に保持し、それより古いパケットは待ち行列から脱落させるものである。

【0035】このシステムの物理リンク層の重要な点は、例えばパケットが送信された後、いつ逆方向リンク層がフリーになるかを決定することである。特定のシステム要件例えばADSL、HFC、等によれば占有性（使用されたバンド幅）を確認する公知のハードウェアが存在し、そして順方向リンクと逆方向リンクの他の状態パラメータが存在する。

【0036】逆方向リンクがフリーで送信できると一旦決定されると、スケジューラ300は伝送管理ポリシーに従って、送信されるべき特定の待ち行列151a, . . . , 151nからパケットを選択することにより待ち行列にサービスする。このような伝送管理方針は、公平重み付き待ち行列化を含み、そこでは各待ち行列は、等しいサービス時間が与えられるというものである。

【0037】このため重み付きスキームに従ってACKパケットを伝送する場合には、プロセッサは順方向（下流方向）リンク90に関し限界パラメータ情報を決定し、これは例えば下流方向伝送レート C_{down} （ビット/秒）、下流方向（順方向リンク）ボトルネックバッファ B_{down} のサイズと、さらに例えば上流方向伝送レート C_{up} （ビット/秒）と、ACKパケットの長さ L_{ACK} の情報を決定し、そしてスケジューラは特定のACKパケットに対する重み W_{ACK} を計算し、その重みに従ってACKパケットを転送して伝送バンド幅を最大にする。

【0038】このパラメータ情報は、ネットワーク接続の往復遅延時間、順方向バッファサイズ、正規化された非対称性を表す。逆方向（上流方向）リンクを介した伝送用データパケットの選択も同様に決定できる。

【0039】本発明の理解を助けるために、図4を説明する。同図はプロセッサ40とウェブサイト10との間に形成されたPC50の一例を示し、ウェブサイト10からプロセッサ40へデータネットワーク20を介して情報をダウンロードするためのものである。このPC50は、ヘッドエンドプロセッサ53と下流方向リンク35と複数の加入者を接続する従来のケーブルリンクとを含む。

【0040】下流方向リンク35は、プロセッサ40を含む複数の加入者にそれぞれのインタフェース装置（図示せず）例えばケーブルモデムを介して接続される。プロセッサ40がパソコンの場合には、接続はユーザのリクエストにより従来方法により開始される。これは通常公知のTCPプロトコルを用い、そしてメッセージを上流方向リンク36例えば電話線を介してケーブルヘッドエンドプロセッサ53に送信することにより行われる。

【0041】その後ケーブルヘッドエンドプロセッサ53は、このメッセージをデータネットワーク20を介してメッセージ内に特定されたインターネットウェブサイト（10）に転送する。そして次にこのウェブサイト10は、要求された情報をデータパケットの形で下流方向リンクを介してプロセッサ40に送信する。フロントエンドプロセッサ41は、下流方向パケットを受信し、それを適当な処理装置42に転送する。そしてこの処理は、従来のACKパケットをTCPフォーマットの形で生成してこのパケットをACK待ち行列43内に記憶してさらに上流方向リンクを介してウェブサイト10に送信する。

【0042】前述したように、ある条件下ではACK待ち行列43は輻輳（満杯）になることがある。この場合、新たなACKは通常廃棄されるが、その理由は待ち行列にロードすることができないからである。この新たなACKは、ACKが既に本発明により待ち行列内に記憶されACKよりも多くの情報を含むので、この新たなACK（即ち、生成された最後のACK）は、廃棄されずに保存される。かくしてACK待ち行列43内に既に登録されたACKは廃棄されて最後のACK用に待ち行列内に空きを形成する。本発明の一実施例によれば、待ち行列内の最も古いACKは廃棄される。

【0043】このACKパケットは、TCP内で別の目的も果たす、即ち、データパケットはデータソースへの受領確認パケットの到達に基づいて制御されたレートでもって送信される。過剰の受領確認パケットが失われると、前述したようにエラー再生情報を保存している場合でも、このフロー制御機構の破壊を引き起こす。例えばN個の受領確認が失われると、(N+1)番目の受領確認の到着によりソースは同時に(N+1)個のパケットを送信する。

【0044】データバッファがこのバーストを吸収できるほど大きくない場合には、データパケットは失われ、これによりTCPはそのウィンドウをドロップして結果的にスループットが低くなる。受領確認バッファの管理は、大多数の受領確認が通過するように行われ、かつ最も新しい受領確認を保存しながら大多数の受領確認を落とすようにして行われる。これはリターンパスに対しては先頭から落とすバッファ理論（drop-from-front buffer discipline）となり、これにより最も新しいエラー再生情報を保存し、かつ最大数のフロー制御情報が可能になる。

【0045】図4から分かるようにプロセッサ40はパケット待ち行列44を有し、このパケット待ち行列44は上流方向リンク36を介して伝送する従来のデータパケットを生成するある逆方向接続（reverse connection）の処理プロセスに関連づけられる。前述したようにある条件下では、逆方向接続のスループットは、上流方向リンクのリソース（資源）が制御された方法で割り当

てられない限りゼロに近づく。

【0046】またこのような割当は本発明によれば、それぞれのサービス、接続、ACKS、データパケット等の専用待ち行列によりかつ重み付けスケジューリング系に従って待ち行列をサービスすることにより行われる。具体的に説明すると、本発明によれば、下流方向リンク35は C_{down} ビット/秒のレートをも有し、上流方向リンク36は C_{up} (ビット/秒) のレートを有すると仮定する。

【0047】さらにTCP接続はいずれかのリンクでサポートされ、かつデータパケットの長さは L_{DATA} で、ACKの長さは L_{ACK} と仮定する。順方向接続に対してはデータパケットに対し、利用可能な最大バンド幅は μ_f

$$\frac{W_{ACK}}{W_{ACK} + W_{data}} C_{up}$$

【0049】本発明によるこれらの待ち行列の最適サービス (スループット) を達成するために、(2) 式の関係は重みとACK待ち行列43との関連を規定し、

$$W_{ACK} \leq \frac{C_{down} L_{ACK}}{B_{down} C_{up}}$$

$$W_{DATA} = 1 - W_{ACK}$$

【0050】(2) 式の W_{ACK} の値が1以上の場合には、下流方向リンクを完全に利用することは不可能である。さらにまた複数のソースがデータパケットを上流方向リンクに供給できることが分かる。この場合、各ソースは、それ自身のパケット/データ待ち行列に関連づけられ、パケットの最も効率的なサービスと上流方向リンクのスループットが達成できる。

【0051】上記の説明は、各接続に対し、利用可能なパケットサービスレートと利用可能なACKサービスレートの比率が順方向パスの利用可能なバッファサイズよりも小さいという条件を満たす重み付きスケジューリングを実行する様々な方法の一実施例を示したものである。即ち各接続に対し、パケットのバーストの平均サイズに等しい $k = \mu_f / \mu_r$ が B_f よりも小さいと、パケットが不必要に失われる。スケジューラ (ラインブロックの先頭) の不完全性に起因して、 k が時間と共に局所的に変化する場合には、最悪のケースを考慮しなければならない。かくして例えば流体モデルスケジューラのようなスケジューラは、割り当てられたレートからの局部変動を最少にする。

【0052】

【発明の効果】以上説明したように本発明は、逆方向リンク上の輻輳点で例えばバッファで以前に受信したACKの代わりに最後のACKを保存し、十分に管理された方法で関連メモリ (バッファ) を含む上流方向リンクの

$= C_{down} / L_{DATA}$ パケット/秒であり、ACKパケットに対し利用可能な最大バンド幅は $\mu_r = C_{up} / L_{ACK}$ である。

【0048】順方向パスと逆方向パスとの間の正規化された非対称 (normalized asymmetry) は、 $k = \mu_f / \mu_r$ で定義される、そしてこれは1以上の値を持つものとする。下流方向リンク35を介して送信されるデータパケットを記憶するため利用可能なバッファのサイズは B_{down} ビットと仮定すると、これは $B_f = B_{down} / L_{DATA}$ パケットに変換される。ACK待ち行列43にサービスするのに、割り当てられた全時間の割合は次式である。

【数1】

(1)

(3) 式の関係は重みとパケット待ち行列44との関係を規定する。

【数2】

(2)

(3)

バンド幅を割り当て、重み付けが下流方向リンクに関連したパラメータに基づいて行われるよう上流方向リンクのバンド幅を割り当てるように、公平待ち行列系あるいは重み付けラウンドロビン系を用いることである。

【図面の簡単な説明】

【図1】従来の単一TCP/インターネット接続を表すブロック図

【図2】単一TCP接続の上流方向リンク上の複合待ち行列バッファを組み込んだブロック図

【図3】本発明の複合待ち行列の要素を表すブロック図

【図4】本発明のデータフロー制御装置を用いたインターネット接続の例を表す図

【符号の説明】

10 ウェブサイト

15 送信者

16 データ端末

20 データネットワーク

25 高速下流方向リンク

26 低速上流方向リンク

30 バッファ

35 下流方向リンク

36 上流方向リンク

40 プロセッサ

41 フロントエンドプロセッサ

42 処理装置

13

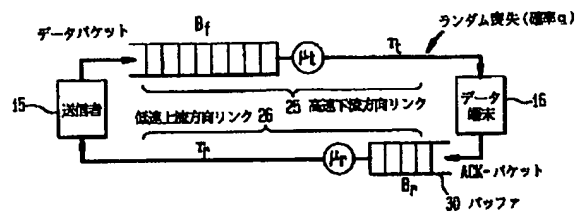
14

43 ACK待ち行列
 44 パケット待ち行列
 50 PC
 53 ヘッドエンドプロセッサ
 55 LAN
 80 WAN/インターネット
 82 サーバ
 85 サービスプロバイダネットワーク端末
 90 順方向（下流方向）リンク

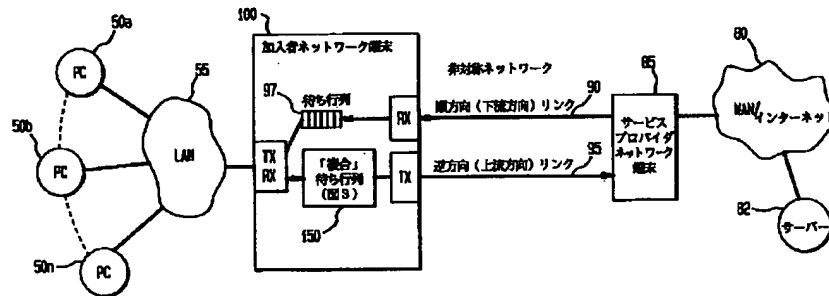
95 逆方向（上流方向）リンク
 97 待ち行列
 100 加入者ネットワーク端末
 150 複合待ち行列
 151 接続毎の待ち行列
 160 メモリと待ち行列マネージャ
 200 クラシーファイヤ
 300 スケジューラ

【図1】

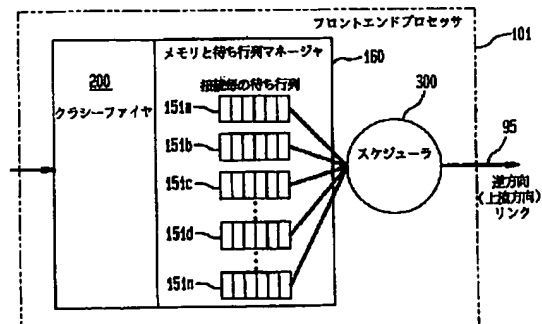
(従来技術)



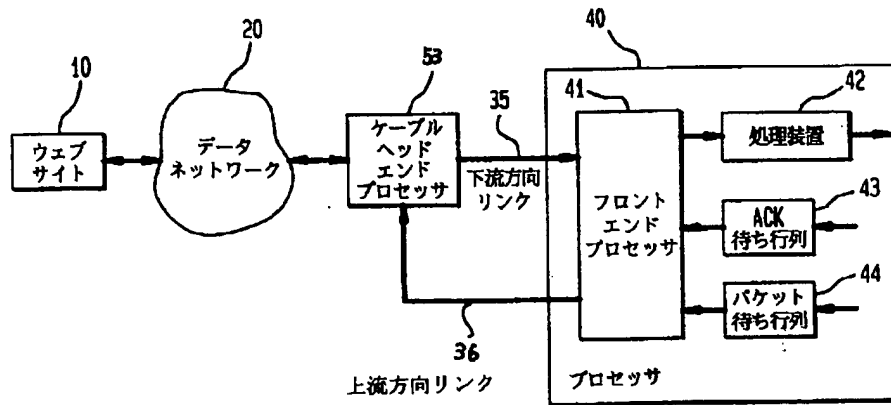
【図2】



【図3】



【図4】



フロントページの続き

(71)出願人 596077259
600 Mountain Avenue,
Murray Hill, New Je
rsey 07974-0636 U. S. A.

(72)発明者 ウバマニュ マドハウ
アメリカ合衆国、61801 イリノイ、アー
バナ、コンベス ストリート 2509

(72)発明者 バーンハード シュター
アメリカ合衆国、07747 ニュージャージー、
アバーディーン、アイダホ レーン

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.